CLAIMS

1. A process for a determining server performance metrics in a network, comprising the steps of:

providing service metric probe means resident on a server for determining the service availability and metric measurements of types of services provided by a content delivery machine;

providing latency probe means resident on a server for determining the latency of various servers within said network;

wherein said service metric probe means consults a configuration file containing each DNS name in its area and the set of services associated with each DNS name;

wherein said services include, but are not limited to: HTTP, HTTPS, FTP, streaming media, and/or generic SNMP; and

wherein said latency probe means calculates the latency from its location to a client's location.

2. The process of claim 1, wherein each server in said network has a metric test associated with each service supported by said server.

3. The process of claim 1, wherein said service metric probe means periodically performs metric tests on the servers within its area, and wherein said service metric probe means records the metric results from said periodic tests.

4. The process of claim 1, wherein said latency probe means calculates the round trip time for sending a packet to a client to obtain the latency value, and wherein the round trip time tests that said latency probe means performs, includes, but are not limited to: PING, UDP Reverse Name lookup, and/or UDP Packets to high number ports.

5. The process of claim 1, wherein when said latency probe means sends a UDP Packet probe to high number ports that fails, said latency probe means resends said UDP Packet probe with a low TTL number and increments the TTL until failure occurs, the last successful TTL value will indicate the partial latency data.

6. The process of claim 1, further comprising the step of:

providing at least one DNS server.

7. The process of claim 6, wherein said service metric probe means sends an update to all of said DNS servers in said network that consists of all tests since the last update.

8. The process of claim 6, wherein said latency probe means updates said DNS servers with the clients' latency data.

9. The process of claim 6, wherein a DNS server uses said latency test data updates to determine the closest server to a client.

10. The process of claim 6, wherein a DNS server uses said test result updates to determine the best server to return for a given DNS name.

11. The process of claim 6, wherein said service metric probe means sends a packet request to a server and receives, in response, a packet containing the various metrics of the server, and wherein said service metric probe means combines the server metrics to arrive at a load metric which is sent to said DNS servers.

12. A process for a determining server performance metrics in a network, comprising the steps of:

providing service metric probe means resident on a server for determining the service availability and metric measurements of types of services provided by a content delivery machine;

providing latency probe means resident on a server for determining the latency of various servers within said network;

providing at least one DNS server;

wherein said service metric probe means sends an update to all of said DNS servers in said network that consists of all service availability and metric measurements since the last update; and

wherein said latency probe means updates said DNS servers with clients' latency data.

13. The process of claim 12, wherein said service metric probe means consults a configuration file containing each DNS name in its area and the set of services associated with each DNS name, and wherein said services include, but are not limited to: HTTP, HTTPS, FTP, streaming media, and/or generic SNMP.

14. The process of claim 12, wherein said latency probe means calculates the latency from its location to a client's location.

15. The process of claim 12, wherein each server in said network has a metric test associated with each service supported by said server.

16. The process of claim 12, wherein said service metric probe means periodically performs metric tests on the servers within its area, and wherein said service metric probe means records the metric results from said periodic tests.

17. The process of claim 12, wherein said latency probe means calculates the round trip time for sending a packet to a client to obtain the latency value, and wherein the round trip time tests that said latency probe means performs, includes, but are not limited to: PING, UDP Reverse Name lookup, and/or UDP Packets to high number ports.

18. The process of claim 12, wherein when said latency probe means sends a UDP Packet probe to high number ports that fails, said latency probe means resends said UDP Packet probe with a low TTL number and increments the TTL until failure occurs, the last successful TTL value will indicate the partial latency data.

19. The process of claim 12, wherein a DNS server uses said latency test data updates to determine the closest server to a client.

20. The process of claim 12, wherein a DNS server uses said test result updates to determine the best server to return for a given DNS name.

21. The process of claim 12, wherein said service metric probe means sends a packet request to a server and receives, in response, a packet containing the various metrics of the server, and wherein said service metric probe means combines the server metrics to arrive at a load metric which is sent to said DNS servers.

22. A program storage medium readable by a computer, tangibly embodying a program of instructions executable by the computer to perform method steps for a determining server performance metrics in a network, comprising the steps of:

providing service metric probe means resident on a server for determining the service availability and metric measurements of types of services provided by a content delivery machine;

providing latency probe means resident on a server for determining the latency of various servers within said network;

wherein said service metric probe means consults a configuration file containing each DNS name in its area and the set of services associated with each DNS name;

wherein said services include, but are not limited to: HTTP, HTTPS, FTP, streaming media, and/or generic SNMP; and

wherein said latency probe means calculates the latency from its location to a client's location.

23. The method of claim 22, wherein each server in said network has a metric test associated with each service supported by said server.

24. The method of claim 22, wherein said service metric probe means periodically performs metric tests on the servers within its area, and wherein said service metric probe means records the metric results from said periodic tests.

25. The method of claim 22, wherein said latency probe means calculates the round trip time for sending a packet to a client to obtain the latency value, and wherein the round trip time tests that said latency probe means performs, includes, but are not limited to: PING, UDP Reverse Name lookup, and/or UDP Packets to high number ports.

26. The method of claim 22, wherein when said latency probe means sends a UDP Packet probe to high number ports that fails, said latency probe means resends said UDP Packet probe with a low TTL number and increments the TTL until failure occurs, the last successful TTL value will indicate the partial latency data.

27. The method of claim 22, further comprising the step of:

providing at least one DNS server.

28. The method of claim 27, wherein said service metric probe means sends an update to all of said DNS servers in said network that consists of all tests since the last update.

29. The method of claim 27, wherein said latency probe means updates said DNS servers with the clients' latency data.

30. The method of claim 27, wherein a DNS server uses said latency test data updates to determine the closest server to a client.

31. The method of claim 27, wherein a DNS server uses said test result updates to determine the best server to return for a given DNS name.

32. The method of claim 27, wherein said service metric probe means sends a packet request to a server and receives, in response, a packet containing the various metrics of the server, and wherein said service metric probe means combines the server metrics to arrive at a load metric which is sent to said DNS servers.

33. A program storage medium readable by a computer, tangibly embodying a program of instructions executable by the computer to perform method steps for a determining server performance metrics in a network, comprising the steps of:

providing service metric probe means resident on a server for determining the service availability and metric measurements of types of services provided by a content delivery machine;

providing latency probe means resident on a server for determining the latency of various servers within said network;

providing at least one DNS server;

wherein said service metric probe means sends an update to all of said DNS servers in said network that consists of all service availability and metric measurements since the last update; and

wherein said latency probe means updates said DNS servers with clients' latency data.

34. The method of claim 33, wherein said service metric probe means consults a configuration file containing each DNS name in its area and the set of services associated with each DNS name, and wherein said services include, but are not limited to: HTTP, HTTPS, FTP, streaming media, and/or generic SNMP.

35. The method of claim 33, wherein said latency probe means calculates the latency from its location to a client's location.

36. The method of claim 33, wherein each server in said network has a metric test associated with each service supported by said server.

37. The method of claim 33, wherein said service metric probe means periodically performs metric tests on the servers within its area, and wherein said service metric probe means records the metric results from said periodic tests.

38. The method of claim 33, wherein said latency probe means calculates the round trip time for sending a packet to a client to obtain the latency value, and wherein the round trip time tests that said latency probe means performs, includes, but are not limited to: PING, UDP Reverse Name lookup, and/or UDP Packets to high number ports.

39. The method of claim 33, wherein when said latency probe means sends a UDP Packet probe to high number ports that fails, said latency probe means resends said UDP Packet probe with a low TTL number and increments the TTL until failure occurs, the last successful TTL value will indicate the partial latency data.

40. The method of claim 33, wherein a DNS server uses said latency test data updates to determine the closest server to a client.

41. The method of claim 33, wherein a DNS server uses said test result updates to determine the best server to return for a given DNS name.

42. The method of claim 33, wherein said service metric probe means sends a packet request to a server and receives, in response, a packet containing the various metrics of the server, and wherein said service metric probe means combines the server metrics to arrive at a load metric which is sent to said DNS servers.